MODFLOW – Automated Parameter Estimation

1 Introduction

The *MODFLOW-Model Calibration* tutorial describes the basic calibration tools provided in GMS. It illustrates how head levels from observation wells and observed flows from streams can be entered into GMS and how these data can be compared to model computed values. It also describes how a trial and error method can be used to iteratively adjust model parameters until the model computed values match the field observed values to an acceptable level of agreement.

In many cases, calibration can be achieved much more rapidly with an inverse model. An inverse model is a utility that automates the parameter estimation process. The inverse model systematically adjusts a user-defined set of input parameters until the difference between the computed and observed values is minimized.

GMS contains an interface to an inverse model called PEST and this tutorial illustrates how to calibrate a MODFLOW model using PEST. Since this tutorial assumes you understand how to enter field observed values, you should complete the *MODFLOW-Model Calibration* tutorial, prior to beginning this tutorial.

1.1 Contents

1.2 Outline

This is what you will do:

1. Open a MODFLOW model and solution.
2. Define conditions.
3. Run PEST.
4. Load optimal parameter values.
5. Create pilot points and define other conditions for another PEST simulation.
6. Run PEST and view results.

1.3 Required Modules/Interfaces

You will need the following components enabled to complete this tutorial:

- Grid
- Geostatistics
- Map

- MODFLOW
- Inverse Modeling

You can see if these components are enabled by selecting the *File* | *Register* command.

2 Description of Problem

The model we will be calibrating in this tutorial is the same model featured in the *MODFLOW-Model Calibration* tutorial. The model includes observed flow data for the stream and observed heads at a set of scattered observation wells. The conceptual model for the site consists of a set of recharge and hydraulic conductivity zones. These zones will be marked as parameters and an inverse model will be used to find a set of recharge and hydraulic conductivity values that minimize the calibration error. We will utilize two methods for model parameterization: polygonal zones and pilot point interpolation.

3 Getting Started

Let's get started.

1. If necessary, launch GMS. If GMS is already running, select the *File* | *New* command to ensure that the program settings are restored to their default state.

4 Reading in the Project

First, we will read in the modeling project:

1. Select the *Open* button.
2. Locate and open the **tutfiles\MODFLOW\inverse** directory.
3. Open the file entitled **bigval.gpr**.

You should see a MODFLOW model with a solution and a set of map coverages. Three of the coverages are the source/sink, recharge, and hydraulic conductivity coverages used to define the conceptual model. The active coverage contains a set of observed head values from observation wells. If you switch to the source/sink coverage, you will notice that an observed flow value has been assigned to the stream network.

5 Model Parameterization

The first step in setting up the inverse model is to "parameterize" the input. This involves identifying which parts of the model input we want the inverse model utility to optimize. We will utilize two approaches for parameterization. In the first attempt at parameter estimation, we will use the zonal approach. This involves identifying polygonal zones of hydraulic conductivity and recharge, marking the zones as parameters, and assigning a starting value for each zone. PEST will then adjust the K or recharge values assigned to the zones as it attempts to minimize the residual error between computed vs. observed heads and flows. In the second part of this tutorial, we will use the pilot point method in conjunction with PEST to parameterize hydraulic conductivity. With the pilot point method, we define a set of scatter points where the hydraulic conductivity is assigned. Each point acts as an independent parameter and the K values for the grid are interpolated from the pilot points. The pilot point method allows for a more continuous and potentially more complex distribution of values throughout the model domain. It also alleviates the modeler from having to define the distribution of the zones, a process that can be difficult given limited data and is often done in an arbitrary manner.

6 Defining the Parameter Zones

For our first attempt at parameter estimation, we will define a set of parameter zones. The conceptual model approach utilized in GMS is ideally suited for this task since the conceptual model consists of recharge and K zones defined with polygons. We will mark the polygons as parameter zones by assigning a “key value” to each polygon. The key value should be a value that is not expected to occur elsewhere in the MODLOW input file. A negative value typically works well.

We will use seven parameter zones consisting of four hydraulic conductivity zones and three recharge zones. The number of observations is eleven, consisting of ten observation wells and one stream flow value. Note that the number of parameters being estimated should always be less than the number of observations when you are using parameter zones.

6.1 Setting up the Hydraulic Conductivity Zones

First we will set up the hydraulic conductivity zones. The hydraulic conductivity polygons are shown in Figure 1. The five polygons will be used to define four parameter zones. The key values associated with each of the four zones are shown on the polygons in the figure.

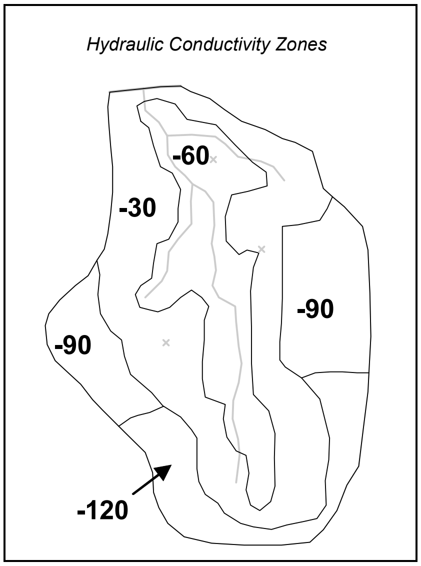

Figure 1. Hydraulic Conductivity Zones and Parameter Key Values.

To assign the key values to the polygons:

1. Expand the *BigVal* conceptual model by clicking the plus symbol next to it in the *Project Explorer*.
2. Switch to the *Hydraulic Conductivity* coverage by selecting it from the *Project Explorer*.
3. Select the *Select Polygons* tool.
4. Double click on each of the polygons shown in Figure 1 and assign the appropriate key value to the *Horizontal K* input field.

6.2 Setting up the Recharge Zones

Next, we will set up the recharge zones. The recharge polygons are shown in Figure 2. The five polygons will be used to define four parameter zones. The polygon at the top end of the model is relatively small, it is isolated from the majority of the observation wells, and it is down gradient from the wells. As a result, it is not a good candidate for parameter estimation. We will fix the recharge in this zone at zero. The key values will be associated with the other four polygons to define three parameter zones as shown.

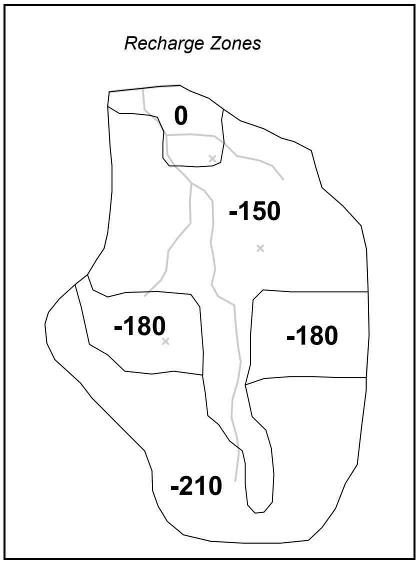

Figure 2. Recharge Zones and Parameter Key Values.

To assign the key values to the polygons:

1. Switch to the *Recharge* coverage by selecting it from the *Project Explorer*.
2. Double click on each of the polygons shown in the previous figure and assign the appropriate key value to the *Recharge rate* input field.

6.3 Mapping the Key Values to the Grid Cells

Once the key values are assigned to the polygons, they must be mapped to the cells in the MODFLOW grid.

3. Select the *Feature Objects | Map → MODFLOW* command.
4. Select *OK* at the prompt.

7 Selecting the Parameter Estimation Option

Before we edit the parameter data, we will turn on the *Parameter Estimation* option. This option is turned on in the *Global Options* dialog.

1. Select the *3D Grid Data* folder in the *Project Explorer*.
2. Select the *MODFLOW | Global Options* command.
3. Select the *Parameter Estimation* option in the *Run options* section of the dialog.

8 Starting Head

The head contours currently displayed on the grid are from a forward run of a MODFLOW simulation using the starting parameter values. Before running PEST, we will copy the computed heads to the Starting Heads array. This will ensure that each time PEST runs MODFLOW, the starting head values will be reasonably close to the final head values and MODFLOW should converge quickly.

1. Select the *Starting Heads* button.

2. Select the *3D Data Set → Grid* command.

3. Expand the *bigval (MODFLOW)* folder if necessary.

4. Select the **Head** data set under the *bigval (MODFLOW)* folder and click *OK*.

5. Select the *OK* button to exit the *Starting Head* dialog.

6. Select the *OK* button to exit the *Global Options* dialog.

9 Editing the Parameter Data

Next, we will create a list of parameters and enter a starting, minimum, and maximum value for each.

1. Select the *MODFLOW* | *Parameters* command.

The *Parameters* section of this dialog is used to manage a list of the parameters used by the inverse model. The *New* button can be used to create a set of parameters one at a time. Each parameter has several properties, including a name, a key value, a type, a starting value, a minimum value, a maximum value, a usage field, and a log-transform field. Rather than creating each parameter one at a time, in most cases the parameter list can be automatically generated by GMS using the *Initialize From Model* button. When this button is selected, GMS searches the MODFLOW input data for likely key values and creates a list of parameters. If a parameter is not automatically found by GMS, it can be added using the *New Parameter* button.

2. Select the Initialize from Model button.

Note that all seven parameters were automatically found. Also note that the parameters have been given a default name. The next step is to enter a starting, minimum, and maximum value for each parameter. Special care should be taken when selecting the starting values. In most cases, using arbitrary starting values will cause the inverse model to fail to converge. The closer the starting values are to the final parameter values, the greater the chance that the inverse model will converge. The starting values we will use were found after a few iterations of manual (trial and error) calibration.

3. For each parameter check the *Param. Est. Solve* toggle.

4. Enter the following data into the parameters spreadsheet:

Name	Value	Min Value	Max Value
HK_30	1.2	0.003	30
HK_60	2.4	0.003	30
HK_90	0.6	0.003	30
HK_120	0.2	0.003	30
RCH_150	0.0001	1e-10	0.0001
RCH_180	0.00008	1e-10	0.0001
RCH_210	0.00006	1e-10	0.0001

5. Ensure that the *Log Xform* option is turned on for each of the hydraulic conductivity parameters.

6. Select *OK* to exit the dialog.

10 Max. Iterations

Finally, we will increase the maximum number of iterations used by the solver package. This will increase the likelihood that MODFLOW will converge at each iteration.

1. Select the *MODFLOW | PCG2 Package* command.

2. Change the *Maximum outer iterations* to **100** and the *Maximum inner iterations* to **100**.

3. Select the *OK* button.

11 Saving the Project and Running PEST

We are now ready to save the project and run PEST.

1. Select the *File | Save As* command.

2. Save the project with the name **mfpest_zones.gpr**.

3. Select the *MODFLOW | Run MODFLOW* command.

PEST is now running. The spreadsheet in the top-right-hand corner of the dialog shows the error and the parameter values for each iteration of the parameter estimation process. The plot on the left shows the error for each iteration. When PEST is finished running you can view the optimum parameter values in the spreadsheet. Once PEST has found the optimum parameter values MODFLOW will run a forward run with the optimum values and output the head solution. This will be the solution that we read in.

Note: If you have room on your screen, you may wish to resize the output window by dragging the handle in the lower right corner of the window.

12 Viewing the Solution

Once MODFLOW is done running you can read in the solution.

1. Make sure that the *Read solution on exit* toggle is checked and select the *Close* button.

The contours currently shown on the 3D grid are the heads from the MODFLOW run with the optimum parameter values. We will now look at the observation targets in the map module and the error associated with this model run.

2. Select the *Observation Wells* coverage. Notice that the residual error has been greatly reduced for all monitoring wells.
3. Select the *Sources & Sinks* coverage from the *Project Explorer*. Notice that the observation target (located near the top of the stream) shows very little residual error.
4. Select the *Select Arc Group* tool.
5. Select the arc group by clicking on a river arc. Notice in the edit strip at the bottom of the graphics window the computed and observed flow is reported.

Next, we will look at some global error norms.

6. Right-click on the **mfpest_zones (MODFLOW)** solution in the *Project Explorer* and select the *Properties* command.

This command brings up a spreadsheet showing the residual error (computed – observed) from this model run. The spreadsheet shows the error from the head observations, the flow observations, and the combined weighted observations. You may wish to compare the values from this run to the *bigval* MODFLOW solution.

7. Select the *Done* button to exit the dialog.

13 Loading Optimal Parameter Values

When you are finished using the inverse model it is often desirable to load and view the optimal parameter values. We will now load the optimal parameters.

1. Select the *3D Grid Data* folder in the *Project Explorer*.
2. Select the *MODFLOW | Parameters* command.
3. Select the *Import Optimal Values* button.
4. Select the **mfpest_zones.par** file and select the *Open* button.

Notice that the starting value for all of the parameters has now been changed to the optimal value computed by the inverse model.

5. Select the *OK* button to exit the dialog.

14 Viewing Parameter Sensitivities

In addition to computing the optimal parameter values, PEST will compute the sensitivities of each parameter. This information is printed to the *.sen file at each parameter estimation iteration. GMS will create a plot of this information using the plot wizard.

1. Select the *Plot Wizard* button.
2. Select the *Parameter Sensitivities* plot.
3. Select the *Next* button.
4. Select the *Finish* button.

Your plot should be similar to the one shown below.

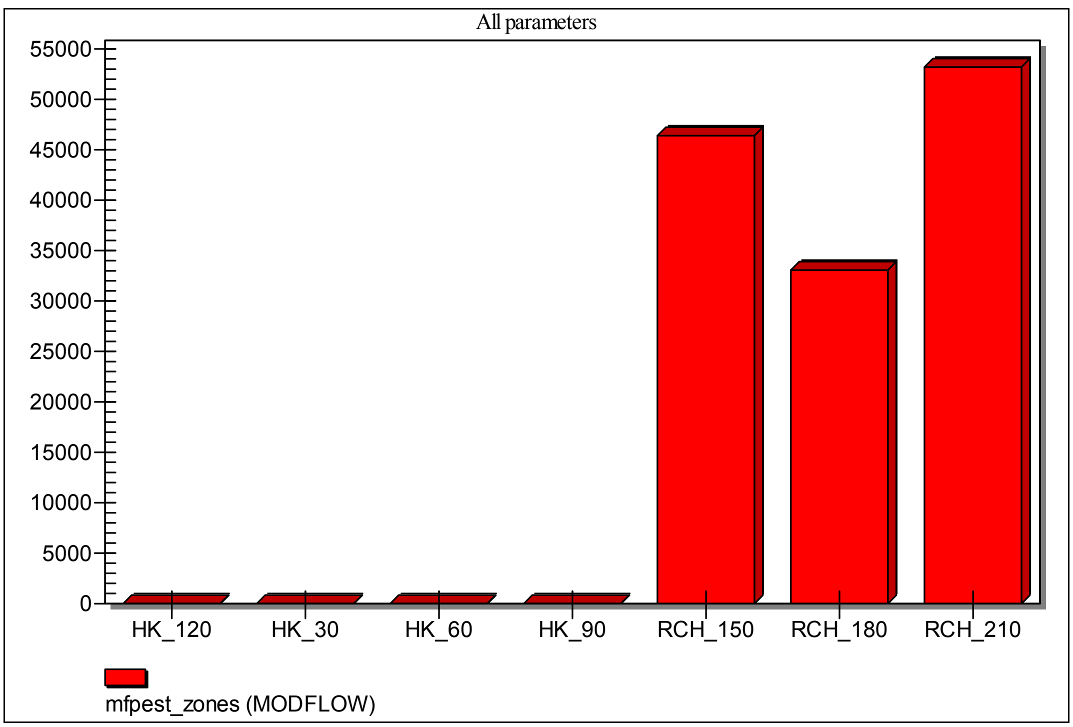

Figure 3 Parameter Sensitivity Plot

Notice that the sensitivity for the ***RCH_*** parameters is much greater than the ***HK_*** parameters. For a more in depth explanation of the calculations used to create this file, the user is referred to the PEST documentation (section 5.3.2 of the PEST Manual). The parameter sensitivity information is useful in identifying the parameters that have the greatest effect on the model and those parameters that have a little or no effect on the model. The user may wish to remove or hold constant the "insensitive" parameters in another PEST model run.

5. Close the Plot Window and maximize the Graphics Window prior to continuing onto the next step in the tutorial.

15 Pilot Points

Now we will run PEST again to solve the same problem. However, we will use a different method for parameterization. For this problem we will use the zonal approach for recharge, but we will use the pilot point interpolation method for hydraulic conductivity.

Pilot points can be thought of as a 2D scatter point set. Instead of creating a zone and having the inverse model estimate one value for the entire zone, the value of the parameter within the zone is interpolated from the pilot points. Then the inverse model estimates the values at the pilot points. Figure 4 shows a set of pilot points used to estimate horizontal hydraulic conductivity. Notice how the hydraulic conductivity now varies from cell to cell. When the inverse model runs, the values at the pilot points are adjusted and the "surface" defining the variation of K values is warped until the objective function is minimized.

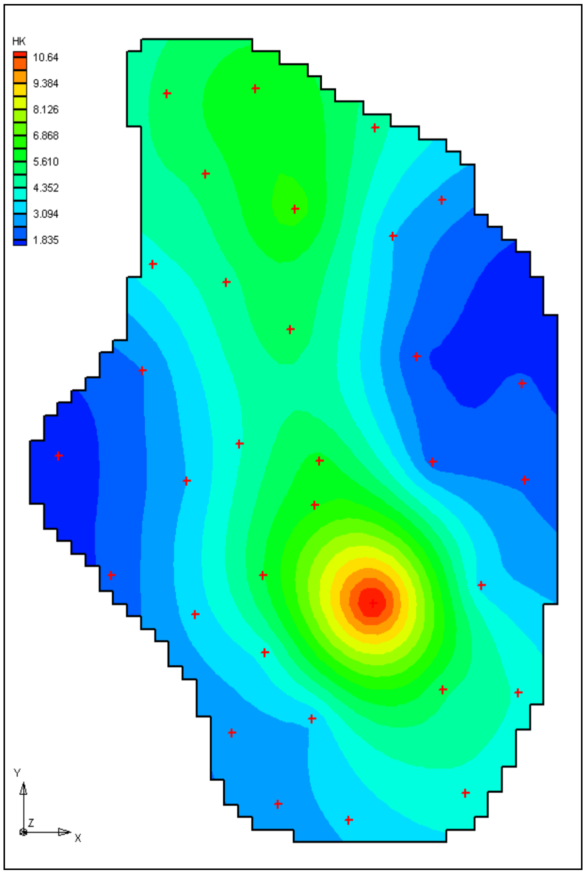

Figure 4. Pilot Points and Resulting Conductivity Field

PEST provides an option for the pilot point method called "regularization". Regularization imposes an additional measure of "stiffness" to the parameter being interpolated via a "homogeneity" constraint. In the absence of any strong influence from the PEST objective function, this constraint causes values at pilot points to approximate the mean value of adjacent pilot points. This constraint makes the inversion process much more stable and makes it possible to violate one of the typical constraints associated with parameter estimation: namely, the requirement that the number of parameters must be less than the number of observations. With regularization, the number of parameters can greatly exceed the number of observations. As a result, complex hydraulic conductivity distributions can be defined, resulting in extremely low residual error. The pilot point method with regularization is an incredibly powerful feature of PEST.

16 Creating Pilot Points

Next, we will create the pilot points that define the hydraulic conductivity distribution for our model. The pilot points are defined as 2D scatter point sets in GMS. We will create about 15 points. In a normal case, we may use 50 or more points. However, additional points slow down the calibration process and 15 points are adequate to illustrate the process with this particular model.

1. In the *Project Explorer* right-click on the empty space and then, from the pop-up menu, select the *New | 2D Scatter Set* command.

2. Right-click on the new 2D scatter set in the *Project Explorer* and select the rename command. Enter **HK** as the name.

3. Select the *2D Scatter Data* folder in the *Project Explorer*.

4. Select the *Scatter Points | Scatter Point Settings* command.

5. Enter **2.0** for the *Default data set value* and select the *OK* button.

6. Select the *Create Scatter Points* tool.

7. Click out a set of scatter points similar to those shown in Figure 5. Create about 15 points. Don't worry about the exact location of the points, as long as they are distributed in a reasonable fashion.

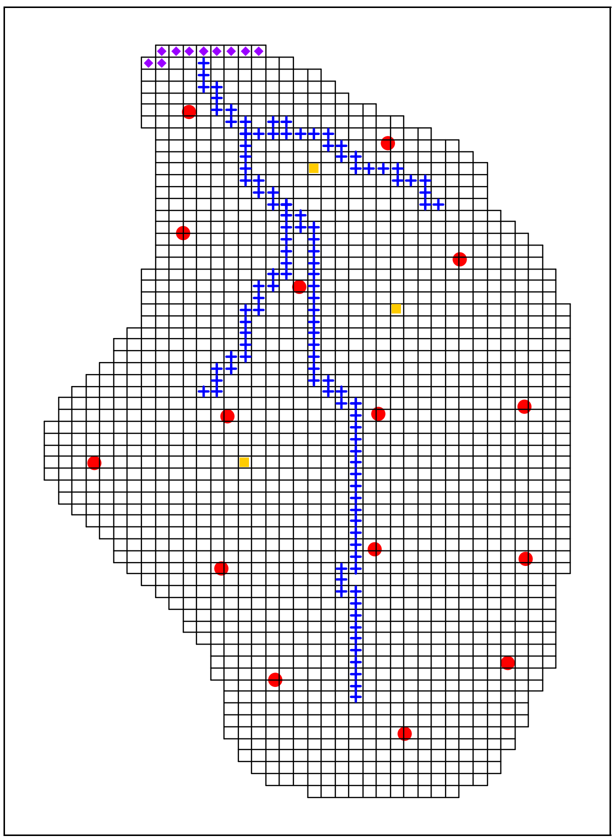

Figure 5. Placement of Scatter Points

17 Entering the HK parameter

In our previous example we had four hydraulic conductivity parameters. For this problem we are estimating the hydraulic conductivity for the entire layer with pilot points. Therefore, we only need one parameter for hydraulic conductivity. This parameter will then be linked to the pilot points using the *Parameters* dialog.

17.1 Creating One Parameter Zone

We will change the values that we have assigned to the polygons so that there is only one hydraulic conductivity parameter.

1. Select the **Hydraulic Conductivity** coverage from the *Project Explorer*.
2. Select the *Select Polygon* tool .
3. Double click on each of the polygons in the coverage and change the value of the *Horizontal K* input field to **–30**.
4. Select the *Feature Objects* | *Map → MODFLOW* command to convert the conceptual model.
5. Select the *All applicable coverages* option and click *OK*.

17.2 Editing the Parameters

Now we will edit the parameters that we currently have defined for MODFLOW.

1. Select the *3D Grid Data* folder in the *Project Explorer*.
2. Select the *MODFLOW* | *Parameters* command.
3. Select the *HK_60* parameter by clicking in the cell where the parameter name *HK_60* is specified.
4. Select the *Delete* button.
5. Repeat this process for the *HK_90* and *HK_120* parameters.
6. Turn on the *Pilot Points* option for parameter *HK_30* by selecting the drop down arrow in the *Value* column. Then select **<Pilot points>** from the drop down list.

The interpolation options associated with the pilot points can be changed by clicking on the small button above the drop down arrow in the *Value* column.

7. Click on the button above the drop down arrow in the *Value* column for parameter *HK_30*. This brings up the interpolation options dialog. Here you can select the scatter point set and data set used with your parameter as well as the interpolation scheme.
8. The defaults are appropriate in this case so we won't change anything. Select the *OK* button twice to exit the dialogs.

17.3 Limiting the Number of Parameter Estimation Runs

In the interest of time we will limit the number of iterations that PEST does for our problem

1. Select the *MODFLOW | Parameter Estimation* command.
2. Change the *Max # of iterations* to **10**.
3. Select *OK* to exit the dialog.

18 Saving the Project and Running PEST

We are now ready to save and run PEST.

1. Select the *File | Save As* command.
2. Save the project with the name **mfpest_pilot.gpr**.
3. Select the *MODFLOW | Run MODFLOW* command.

PEST is now running. This dialog is similar to the one used when MODFLOW is running in Parameter Estimation mode. The error and parameter values are shown in the spreadsheet in the upper right side of the dialog and the plot on the left shows the error. In this case you may notice some strange parameter names like sc1v1. These names were automatically generated and assigned to the scatter points.

PEST will take several minutes to run. You should see the residual error go to an extremely small value. When PEST is finished, you will see a message in the text portion of the window and the *Abort* button will change to *Close*.

19 Viewing the Solution

Once PEST is finished, you can read in the solution.

1. Select the *Close* button. Make sure that the *Read solution on exit* toggle is checked.

The contours currently shown on the 3D grid are the heads from the MODFLOW run with the optimum parameter values. We will now look at the observation targets in the map model and the error associated with this model run.

2. Select the *Sources & Sinks* coverage from the *Project Explorer*. Notice that the observation target on the arc group almost exactly matches.
3. Select the *Select Arc Group* tool.
4. Select the arc group by clicking on the river arc. Notice in the edit strip at the bottom of the graphics window the computed and observed flow is reported.

5. Select the *Observation Wells* coverage from the *Project Explorer*.

6. Right-click on the *pest (MODFLOW)* solution in the *Project Explorer* and select the *Properties* command.

This command brings up a spreadsheet showing the error from this model run. The spreadsheet shows the error from the head observations, the flow observations, and the combined weighted observations. Note that these values are lower than the values we obtained using the zonal parameters approach.

7. When you finish viewing the properties select *Done.*

20 Viewing the Final Hydraulic Conductivity

When PEST ran a new conductivity value was estimated at each of the scatter points used with the **HK_30** parameter. Now we will read in the optimal parameter values as determined by PEST. Reading in the optimal parameter values will create a new data set for our scatter point set. Then we can see our final hydraulic conductivity field.

1. Select the *3D Grid Data* folder in the *Project Explorer*.

2. Select the *MODFLOW | Parameters* command.

3. Click the *Import Optimal Values* button.

4. Select the **mfpest_pilot.par** file and select the *Open* button. Notice that the starting values for the parameters have changed.

5. View the pilot point options by selecting the button above the drop down arrow in the *Value* column for parameter *HK_30*.

Notice the **HK_30 (mfpest_pilot)** data set is listed as the one being used by the pilot points. This data set was imported when you imported the optimal values and represents the optimal values at each pilot point as determined by PEST.

6. Select *OK* to exit both dialogs.

7. Expand the *MODFLOW* item in the *Project Explorer*. Under the *LPF* package select the *HK* data set.

8. Select the *Display | Display Options* command.

9. Turn off the *Contours* option and turn on the *Cell faces* option. Make sure the **Data** colors option is selected below the *Cell faces* option.

10. Select the *Options* button below the *Cell faces* option.

11. In the bottom left portion of the dialog turn on the *Show legend* option.

12. Select *OK* twice to exit both dialogs.

You should now see the final hydraulic conductivity values for your MODFLOW simulation.

21 Conclusion

This concludes the *MODFLOW* - Automated Parameter Estimation tutorial. Here are the things that you should have learned in this tutorial:

- You can use the zonal approach or the pilot point approach with PEST. When using pilot points you may have more parameters than observations because regularization is included in the parameter estimation.

- When PEST finishes, the solution imported into GMS corresponds to the optimal input values. However, the input values in GMS are still the starting values. You must use the *Import Optimal Values* button to replace the starting values with the optimal values.

- When you bring up the *MODFLOW Parameters* dialog, you can change a parameter to use pilot points by using the  button.

- You can view the parameter sensitivities by looking in the *.sen file written by PEST.

- You use 2D scatter points to create pilot points in GMS.

- When you are using pilot points and you select the *Import Optimal Values* button, a new data set is created for the 2D scatter points.

- You can view the final hydraulic conductivity field calculated by PEST by selecting the **HK** item in the *Project Explorer* below the **LPF** package.

